



# The Importance of Preventive Feedback: Inference from Observations of the Stellar Masses and Metallicities of Milky Way Dwarf Galaxies

Yu Lu<sup>1</sup> , Andrew Benson<sup>1</sup> , Andrew Wetzel<sup>1,2,3,10</sup> , Yao-Yuan Mao<sup>4</sup> , Stephanie Tonnesen<sup>1</sup> , Annika H. G. Peter<sup>5,6</sup> , Michael Boylan-Kolchin<sup>7</sup> , and Risa H. Wechsler<sup>8,9</sup>

<sup>1</sup> The Observatories, The Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA

<sup>2</sup> TAPIR, California Institute of Technology, Pasadena, CA 91125, USA

<sup>3</sup> Department of Physics, University of California, Davis, CA 95616, USA

<sup>4</sup> Department of Physics and Astronomy and the Pittsburgh Particle Physics, Astrophysics and Cosmology Center (PITT PACC), University of Pittsburgh, Pittsburgh, PA 15260, USA

<sup>5</sup> CCAPP and Department of Physics, The Ohio State University, 191 W. Woodruff Avenue, Columbus, OH 43210, USA

<sup>6</sup> Department of Astronomy, The Ohio State University, 140 W. 18th Avenue, Columbus, OH 43210, USA

<sup>7</sup> Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712-1205, USA

<sup>8</sup> Kavli Institute for Particle Astrophysics and Cosmology, Department of Physics, Stanford University, Stanford, CA 94305, USA

<sup>9</sup> SLAC National Accelerator Laboratory, Menlo Park, CA, 94025, USA

Received 2017 March 21; revised 2017 August 2; accepted 2017 August 4; published 2017 August 31

## Abstract

Dwarf galaxies are known to have remarkably low star formation efficiency due to strong feedback. Adopting the dwarf galaxies of the Milky Way (MW) as a laboratory, we explore a flexible semi-analytic galaxy formation model to understand how the feedback processes shape the satellite galaxies of the MW. Using Markov Chain Monte Carlo, we exhaustively search a large parameter space of the model and rigorously show that the general wisdom of strong outflows as the primary feedback mechanism cannot simultaneously explain the stellar mass function and the mass–metallicity relation of the MW satellites. An extended model that assumes that a fraction of baryons is prevented from collapsing into low-mass halos in the first place can be accurately constrained to simultaneously reproduce those observations. The inference suggests that two different physical mechanisms are needed to explain the two different data sets. In particular, moderate outflows with weak halo mass dependence are needed to explain the mass–metallicity relation, and prevention of baryons falling into shallow gravitational potentials of low-mass halos (e.g., “pre-heating”) is needed to explain the low stellar mass fraction for a given subhalo mass.

**Key words:** galaxies: abundances – galaxies: dwarf – galaxies: evolution – galaxies: formation – Galaxy: evolution – Galaxy: formation

## 1. Introduction

Our own Galaxy, the Milky Way (MW), provides an excellent laboratory for constraining galaxy formation physics. In particular, the dwarf galaxies in the MW system are excellent places to test feedback models because feedback is expected to be most effective in their shallow gravitational potential wells (e.g., Dekel & Silk 1986; Thoul & Weinberg 1996; Benson et al. 2002a; Okamoto et al. 2010). Observations of kinematic properties of MW satellite galaxies have demonstrated that these dwarf galaxies are dominated by dark matter with mass-to-light ratios as high as  $\sim 100$ – $1000$  (e.g., Mateo 1998; Muñoz et al. 2006; Simon & Geha 2007), which suggests that these satellite galaxies have low baryon fractions. The question is then whether the baryons were lost through strong outflows or instead were never able to condense into those galaxies. In this paper, we attempt to distinguish between these two scenarios using a flexible Semi-Analytic Model (SAM) built on merger trees extracted from high-resolution zoom-in simulations of MW-size halos (Mao et al. 2015). Understanding how feedback works is important not only for understanding galaxy formation but also for constraining the properties of dark matter, as the kinematic effect of the outflow can influence the distribution of dark matter (e.g., Pontzen & Governato 2012; Sawala et al. 2016), which interferes with

observational tests of dark-matter models using astronomical data (e.g., Macciò & Fontanot 2010; Parry et al. 2012; Chau et al. 2016).

Previous theoretical studies have suggested that galaxy formation in small halos can be strongly affected by outflows powered by the injection of supernova energy into the gas (White & Rees 1978; Dekel & Silk 1986) and photoionization heating (Couchman & Rees 1986; Efstathiou 1992; Mo & Miralda-Escude 1996; Thoul & Weinberg 1996; Bullock et al. 2000). Implementing these processes in greater detail in SAMs, many authors have shown that these physical processes are crucial for reproducing the observational properties of dwarf galaxies such as the abundance of the classical dwarfs of the MW system and the properties of low-mass galaxies in the field (e.g., Kauffmann et al. 1993; Benson et al. 2002a; Somerville 2002; Macciò et al. 2010; Guo et al. 2011). More recently, Henriques et al. (2013, 2015) explored the parameter space of the L-GALAXIES SAM and found that when the outflow mass from low-mass halos is kept out of the halo for a long period of time before being reincorporated back into the halo, the model could fit the evolution of the galaxy luminosity function remarkably well. On the other hand, White et al. (2015) found that a redshift-dependent mass-loading factor law can also match the evolution of galaxy stellar mass. Moreover, Hirschmann et al. (2016) tested a number of models with various implementations of feedback, suggesting that the physics of feedback is not well understood.

<sup>10</sup> Caltech-Carnegie Fellow.

Broadly speaking, the feedback processes considered important for the formation of low-mass galaxies can be classified into two categories: (1) *ejective feedback*, which expels baryons from the galaxy to the intergalactic medium (IGM), and (2) *preventive feedback*, which inhibits baryons from accreting onto galaxies or even their dark-matter halos in the first place. In most galaxy formation models, ejective feedback is generally captured by outflows, which remove baryonic mass from the galaxy and deposit it into the IGM (see Benson 2010; Somerville & Davé 2015; and references therein). A common form of preventive feedback considered in SAMs and hydro simulations for low-mass halos is photoionization heating (photoheating for short), which not only reduces radiative cooling but also prevents low-mass halos from accreting their full complement of baryons. Such effects of the reionization of the universe have also been implemented in galaxy formation models by following a “filtering mass” proposed by Gnedin (2000). Indeed, more recent studies have demonstrated that preventive feedback, stronger than what reionization is expected to provide, might also be needed to solve certain problems in galaxy formation (e.g., Mo & Mao 2002; Mo et al. 2005; Lu & Mo 2007; Kauffmann et al. 2013; Lu et al. 2015b; Christensen et al. 2016).

Nevertheless, our understanding of these different forms of feedback is still poor. One can gain deeper understanding of these feedback mechanisms by testing models against observational data. Recently, Hirschmann et al. (2016) implemented various feedback prescriptions into independent SAMs and compared the model predictions of stellar mass, star formation rate, gas fraction, and metallicities of field galaxies with observational data (also see White et al. 2015). Although the authors claimed that none of the tested models appeared to be completely satisfactory in reproducing all of the adopted observational data, they found that the models with strong ejective and preventive feedback were largely degenerate, pointing out the importance of breaking the degeneracy between the two feedback scenarios. Importantly, the authors have found a common feature of the relatively more successful models in their studies, i.e., more than half of the baryons associated with low-mass halos are kept unavailable for cooling or star formation at high redshift. Also, the authors have shown that matching the evolution of the galaxy metallicity–mass relation is generally challenging for galaxy formation models.

It has been shown that the metallicity–luminosity relation of dwarf galaxies provides stringent constraints for models of supernova feedback and reionization (Li et al. 2010; Font et al. 2011; Starkenburg et al. 2013; Gómez et al. 2014; Cousin et al. 2016; Hou et al. 2016). Using an analytic model, Lu et al. (2015a) demonstrated that galaxy metallicities tightly constrain the upper limit for the strength of outflow from a galaxy. If the outflow is too strong, it will blow out too much metal mass to match the observed metallicity–stellar mass relation. Interestingly, Font et al. (2011) found that it was not possible for the GALFORM model (Cole et al. 2000) to simultaneously match the luminosity function and the metallicity–luminosity relation with the default supernova feedback prescription used in previous implementations of the model, in which the efficiency of feedback increases with decreasing halo circular velocity as a steep power law. The authors suggested that the feedback efficiency should saturate for halos with circular velocity  $v_{\text{circ}} \leq 65 \text{ km s}^{-1}$ , and very strong reionization, a combination of cosmic and local reionization with 100% ionizing photon

escape fraction, was needed to simultaneously match the stellar mass function and the mass–metallicity relation of MW satellite galaxies. By varying a few parameters in a SAM, Hou et al. (2014) found that the slope of the metallicity–stellar mass relation is sensitive to the strength of supernova feedback and reionization. Comparing their model results with the observed slope of the metallicity–stellar mass relation, the authors suggested that the local universe was reionized earlier than the cosmic average and a moderate supernova feedback was needed. Using a more sophisticated SAM, Hou et al. (2016) found that, in addition to the early reionization of the local universe, the strength of ejective feedback must evolve with redshift to match the local constraints.

These previous studies provide very useful insight into understanding feedback in galaxy formation. However, these studies were carried out using galaxy formation models with hand-tuned parameters. Owing to the nonlinear nature of galaxy formation, it is often hard to fully capture the impact of multiple processes by varying one parameter at a time (e.g., Henriques et al. 2009; Bower et al. 2010; Lu et al. 2011; Gómez et al. 2012; Benson 2014). A conclusive assessment for a model family can be obtained only if one widely explores the parameter space and studies the collective behavior of an ensemble of models that can match basic observational constraints. In this paper, we employ a Markov Chain Monte Carlo (MCMC) machinery that allows us to explore the parameter space of a flexible galaxy formation model (Lu et al. 2011). We investigate what form of feedback mechanisms are needed to explain two key observational constraints from the MW satellite galaxies: the stellar mass function from the compilation of McConnachie (2012) and the stellar-phase metallicity–stellar mass relation of Kirby et al. (2011). Our goal for this study is to use those two relationships observed for the MW satellite galaxies to break the degeneracy in the model and shed light on understanding the nature of feedback. In particular, we explore an ejective feedback model, in which the effect of feedback is to eject baryons out of halos following star formation, and an extended model, in which a fraction of the baryonic mass is prevented from collapsing into halos in the first place.

The paper is organized as follows. We describe the SAM, the constraining data, and the method of inference in Section 2. The results of the study are presented in Section 3. In detail, we present the posterior distributions of both the ejective and extended models in Section 3.1, show the stellar mass function and stellar-phase metallicity–stellar mass relation predicted by the constrained models and compare them with observations in Section 3.2, and discuss the strength of preventive feedback inferred from the data in Section 3.3. Finally, we summarize the conclusions and discuss the uncertainties of the study in Section 4.

## 2. Methodology

### 2.1. The Merger Trees

In this study, we adopt a suite of dark-matter-only high-resolution zoom-in simulations of MW-size halos, introduced in Mao et al. (2015), to model the stellar mass function and the stellar mass–metallicity relation of MW dwarf galaxies. The final halo virial masses of the simulated halos are in the range of  $M_{\text{vir}} = 10^{12.1 \pm 0.03} M_{\odot}$ , where the virial mass definition follows Bryan & Norman (1998). This halo mass range is

consistent with many observational constraints of the halo mass of the MW (Xue et al. 2008; González et al. 2013; Cautun et al. 2014; Eadie et al. 2015). The mass resolution of the simulations is  $3.0 \times 10^5 h^{-1} M_\odot$  per particle. The softening length in the highest-resolution region is  $170 h^{-1} \text{pc}$  comoving. The cosmological parameters are  $\Omega_M = 0.286$ ,  $\Omega_\Lambda = 0.714$ ,  $h = 0.7$ ,  $\sigma_8 = 0.82$ , and  $n_s = 0.96$ , as described in Mao et al. (2015).

The mass resolution of merger trees extracted from the simulation is sufficient for the model to accurately characterize the properties of galaxies with stellar masses as low as  $\sim 10^4 M_\odot$  (corresponding to halo mass  $\sim 10^9 M_\odot$ ). For more details about this suite of zoom-in simulations, please refer to Mao et al. (2015). We have tested the numerical convergence of the model by artificially reducing the mass resolution and the temporal resolution of the merger tree. We find that reducing the mass resolution by a factor of 10 and the temporal resolution by a factor of 5 produces negligible effects on the predicted stellar mass, metallicity, and star formation histories of the model galaxies.

Lu et al. (2016) find that the whole set of the simulated halos covers a fairly large parameter space in terms of the halo mass-assembly history, concentration, and subhalo mass function. Some host halos do not contain sufficiently high-mass subhalos to host the Magellanic Clouds (MCs) and are therefore inconsistent with the observed stellar mass function, even though the model adopts extreme star formation and feedback parameters to boost the stellar mass to subhalo mass ratio. From the whole set of 46 zoom-in simulation halos, we select 14 whose subhalo population contains at least two subhalos with  $V_{\text{max}} \geq 55 \text{ km s}^{-1}$ , making them close analogues of the MW halo (Lu et al. 2016). The exclusion of other host halos that do not contain high-mass subhalos from this study effectively avoids the situation where the model never fits the observed MW satellite stellar mass function because of the lack of rare high-mass subhalos to host the MCs.

## 2.2. The SAM

In this study, we continue to use the model adopted in Lu et al. (2016), in which we employ flexible parameterizations for the baryonic processes of galaxy formation to encompass a wide range of efficiency for star formation and feedback. The prescriptions for the baryonic processes implemented in the SAM are detailed in Lu et al. (2014b), where we have shown that, aided with MCMC optimization, the model accurately matches the local galaxy stellar mass function and performs well in predicting the stellar mass functions out to  $z \sim 6$ . In this paper, we focus our investigation on the stellar feedback model, which is expected to be the most crucial for shaping low-mass galaxies. Therefore, we only briefly describe the prescriptions that affect the properties of low-mass galaxies in the model.

### 2.2.1. Reionization

Heating due to UV photoionization not only offsets the cooling of halo gas, but also prevents low-mass halos from accreting their full complement of baryons. The suppression of baryonic accretion into halos has been modeled using analytic and numerical models (Gnedin 2000; Hoeft et al. 2006; Okamoto et al. 2008; Noh & McQuinn 2014). Gnedin (2000) showed that the fraction of baryons that can collapse into halos of a given mass in the presence of a photoionizing background can be described in terms of the “filtering mass,”  $M_F$ . Halos less

massive than  $M_F$  accrete less baryonic mass than the universal average. Gnedin (2000) parametrized the collapsed baryon fraction as a function of redshift and halo mass with the expression

$$f_{\text{b, reion}}(z, M_{\text{vir}}) = \frac{f_b}{[1 + 0.26 M_F(z)/M_{\text{vir}}]^3}, \quad (1)$$

where  $f_b$  is the universal baryon fraction and  $M_{\text{vir}}$  is the halo virial mass in dark-matter-only simulations (or equivalently the total mass of dark matter and baryons at the cosmic baryon fraction). The filtering mass is a function of redshift, and this function depends on the reionization history of the universe. Kravtsov et al. (2004) provide a fitting formula for the filtering mass as a function of the redshift at which the first H II regions begin to overlap ( $z_{\text{overlap}}$ ) and the redshift at which most of the medium is reionized ( $z_{\text{reion}}$ ). We use the fitting functions (B2) and (B3) from Appendix B of Kravtsov et al. (2004) to compute the initial fraction of baryons that can collapse as a function of halo mass and redshift,  $f_{\text{b, reion}}$ , with two parameters fixed at  $z_{\text{overlap}} = 11$  and  $z_{\text{reion}} = 10$ . The choice of parameters is consistent with current cosmological constraints (e.g., Komatsu et al. 2009; Planck Collaboration et al. 2016), and the results of our paper are insensitive to the precise values of these parameters.

### 2.2.2. Star Formation

The model assumes that cold gas is distributed in an exponential disk with a scale radius  $r_{\text{gas}}$ , and only gas mass with a surface density higher than a certain threshold,  $\Sigma_{\text{crit}}$ , can form stars (e.g., Kennicutt 1998; Kennicutt et al. 2007; Bigiel et al. 2008). This cold gas mass available for forming stars is then

$$m_{\text{sf}} = M_{\text{cold}} \left\{ 1 - \left[ 1 + \ln \left( \frac{\Sigma_{\text{cold},0}}{\Sigma_{\text{crit}}} \right) \right] \frac{\Sigma_{\text{crit}}}{\Sigma_{\text{cold},0}} \right\}, \quad (2)$$

where  $M_{\text{cold}}$  is the total cold gas mass of the galaxy, and  $\Sigma_{\text{cold},0} = M_{\text{cold}}/(2\pi r_{\text{gas}}^2)$  is the central surface density of the gas disk. This parameterization has two uncertain parameters,  $r_{\text{gas}}$  and  $\Sigma_{\text{crit}}$ , which are clearly degenerate. We rewrite the term  $\Sigma_{\text{cold},0}/\Sigma_{\text{crit}}$  as

$$\frac{\Sigma_{\text{cold},0}}{\Sigma_{\text{crit}}} = \frac{M_{\text{cold}}}{2\pi r_{\text{d},0}^2 \Sigma_{\text{SF}}}, \quad (3)$$

where we define

$$r_{\text{d},0} = \frac{0.035}{\sqrt{2}} R_{\text{vir}}, \quad (4)$$

with  $R_{\text{vir}}$  representing the virial radius of the halo, and the parameter  $\Sigma_{\text{SF}}$  is defined as

$$\Sigma_{\text{SF}} = \left( \frac{r_{\text{gas}}}{r_{\text{d},0}} \right)^2 \Sigma_{\text{crit}}, \quad (5)$$

in units of  $M_\odot \text{pc}^{-2}$ . When we use Equation (3) to replace the term in Equation (2), the uncertainties of two parameters are absorbed by one parameter,  $\Sigma_{\text{SF}}$ .

We assume that the star formation rate is proportional to the mass of star-forming gas and inversely proportional to the



dynamical timescale of the disk,  $\tau_d = r_{d,0}/V_{\text{vir}}$ , yielding

$$\text{SFR} = \alpha_{\text{SF}} \frac{m_{\text{sf}}}{\tau_d}, \quad (6)$$

where  $\alpha_{\text{SF}}$  governs star formation efficiency. It is clear that  $\alpha_{\text{SF}}$  is degenerate with  $\Sigma_{\text{SF}}$  because both of them affect the efficiency of converting gas into stars. We choose to fix  $\alpha_{\text{SF}} = 0.01$ , but allow  $\Sigma_{\text{SF}}$  to vary, so that we do not need to specify two other model parameters,  $\Sigma_{\text{crit}}$  and  $r_{\text{gas}}$ , which are degenerate when the cold gas distribution is not constrained (Lu et al. 2014a).

### 2.2.3. Ejective Feedback

The photoheating effect of reionization is expected to be effective only for halos with a circular velocity  $\lesssim 40 \text{ km s}^{-1}$  (Gnedin 2000; Hoeft et al. 2006; Okamoto et al. 2008). Other feedback processes that are effective on more massive halos are needed to explain the stellar mass function of field galaxies and MW dwarfs. The most widely adopted stellar feedback model is in the form of outflows. Following star formation, the explosion of core-collapse supernovae and stellar winds from high-mass stars provide enough energy to heat up the cold gas in the disk and drive strong outflows to expel baryonic matter from the galaxy. In models, the strength of the outflow is parameterized by a mass-loading factor,  $\eta$ , which characterizes the mass flux expelled out of the galaxy per unit rate of conversion of cold gas mass into stars,  $\dot{M}_{\text{out}} = \eta\phi$ , where  $\dot{M}_{\text{out}}$  is the outflow rate and  $\phi$  is the star formation rate of a galaxy. Regardless of the physics governing the outflow, the outflow mass-loading factor is generally parameterized as a power-law function of the halo circular velocity  $V_c$  as

$$\eta = \alpha_{\text{LD}} \left( \frac{V_c}{220 \text{ km s}^{-1}} \right)^{-\beta_{\text{LD}}}, \quad (7)$$

where  $\alpha_{\text{LD}}$  and  $\beta_{\text{LD}}$  are the model parameters to be constrained. The  $\alpha_{\text{LD}}$  parameter governs the normalization for the mass-loading factor for halos with a circular velocity  $V_c = 220 \text{ km s}^{-1}$ . The parameter,  $\beta_{\text{LD}}$ , controls how rapidly the mass-loading factor varies with the halo circular velocity. In the literature, when  $\beta_{\text{LD}} = 2$ , the corresponding feedback model is referred to as “energy-driven wind,” while when  $\beta_{\text{LD}} = 1$ , it is referred to as “momentum-driven wind.”

The gas mass in the outflow can either be trapped in the host halo or leave the halo depending on the depth of the gravitational potential well and the velocity of the outflow. We introduce two parameters,  $V_{\text{out}}$  and  $\beta_{\text{out}}$ , to describe the behavior. If the circular velocity of the halo is lower than  $V_{\text{out}}$ , the gravitational potential well of the halo is too weak to keep the outflow gas in the halo and most of the outflow mass is expelled from the halo; if the circular velocity of the halo is higher than  $V_{\text{out}}$ , only a smaller fraction of the outflow gas can leave the halo. Quantitatively, the model is parameterized as

$$f_{\text{ej}} = \left[ 1 + \left( \frac{V_c}{V_{\text{out}}} \right)^{\beta_{\text{out}}} \right]^{-1}. \quad (8)$$

In the parameterization, the fraction of the outflow mass that is expelled from the halo decreases with increasing halo circular velocity when  $\beta_{\text{out}} > 0$ . The rest of the outflow gas mass is added into the halo gas, which can cool onto the central galaxy.

In the model, we deposit the expelled mass into a reservoir  $M_{\text{ej}}$  that is tracked for each halo. The mass in this reservoir can increase due to the outflow of the galaxies hosted by the halo and can decrease when the outflow mass is reincorporated back into the halo. A general parameterization for rate of this reincorporation can be written as

$$\dot{M}_{\text{ej}} = -\gamma_{\text{RI}} \frac{M_{\text{ej}}}{\tau_{\text{dyn}}}, \quad (9)$$

where  $\tau_{\text{dyn}}$  is the dynamical timescale of the halo at the virial radius, and  $\gamma_{\text{RI}}$  is a free parameter characterizing the efficiency of the reincorporation. This parameterization is different from the parameterization found in Henriques et al. (2013), where low-mass halos have longer time delays for the ejected mass to be reincorporated back into the halo than high-mass halos, similar to the results obtained by Oppenheimer et al. (2010) using hydrodynamic simulations. However, using recent simulations, Christensen et al. (2016) have shown that the re-accretion times depend only very weakly on halo mass. Anglés-Alcázar et al. (2016) find that re-accretion of outflow mass in high-resolution hydrodynamic simulations occurs over a broad range of timescales. The efficiency parameter,  $\gamma_{\text{RI}}$ , takes into account the deviation from the halo dynamical timescale and the uncertainties. We note that the ejected mass reservoir also tracks the metal mass that is ejected with the outflow. The instantaneous metallicity of the outflow is assumed to be equal to the metallicity of the ISM. The metal mass is assumed to be perfectly mixed in the ejected mass reservoir. We assume perfect mixing as well for metals in other components, including the ISM and the halo gas.

### 2.2.4. Preventive Feedback

In addition to this ejective outflow model, we also hypothesize a preventive feedback model. Differing from the ejective feedback model, the preventive feedback prevents some amount of baryonic mass from being accreted into halos. Such prevention can be due to any feedback process that prevents a halo from accreting its cosmic baryon fraction, for example, strong heating from the local ionizing radiation field in addition to the global ionizing background or from the kinetic energy of strong feedback generated at earlier epochs of the universe that propagates to a large Lagrangian volume. Regardless of the physics behind the prevention, we can capture the effect of preventive feedback using a phenomenological model in which the fraction of baryons that would collapse into a halo is a function of halo mass and redshift. We adopt the following parameterization

$$f_{\text{b,pr}} = \frac{\exp(\gamma_{\text{pr}} z)}{1 + \left( \frac{M_{\text{pr}}}{M_{\text{vir}}} \right)^{\beta_{\text{pr}}}}, \quad (10)$$

where  $\beta_{\text{pr}}$  and  $\gamma_{\text{pr}}$  are the parameters characterizing the mass and redshift dependence, respectively, and  $M_{\text{pr}}$  is a parameter characterizing the mass scale below which the prevention becomes important. We treat  $M_{\text{pr}}$ ,  $\beta_{\text{pr}}$ , and  $\gamma_{\text{pr}}$  as free parameters to be constrained by data. In this model, at a given redshift, the halo baryon fraction decreases with decreasing halo mass as  $f_{\text{b}} \propto M_{\text{vir}}^{\beta_{\text{pr}}}$  when the halo mass is significantly lower than  $M_{\text{pr}}$ . This captures the effect that low-mass halos

form a shallower gravitational potential well, and thus baryon accretion onto low-mass halos can be more strongly affected by any heating processes. For a constant halo mass, the baryon fraction scales with redshift as  $f_b \propto \exp(\gamma_{\text{pr}} z)$ . In the case where  $\gamma > 0$ , the accretion fraction increases with redshift for a constant mass. This captures the effect that the prevention is weaker at early times because the heating is less efficient or that halos at high redshift generally have sufficient gas supply and can accrete baryons at a higher rate. In the model, the total baryonic mass, summing up all of the components in stars, cold gas, and ejected material, is required to be smaller than  $f_b f_{b,\text{reion}} f_{b,\text{pr}} M_{\text{vir}}$ , unless the halo has already converted all its baryonic matter into cold gas or stars. Also, by definition,  $f_{b,\text{pr}}$  cannot be larger than unity. When the numerical value for  $f_{b,\text{pr}}$  exceeds 1 for a halo, it simply means that no prevention is at work, so it is set to unity for the halo.

### 2.2.5. Metallicity

In the model, we follow not only the mass of the baryonic matter in the hot halo, in the disk as cold gas and stars, and in the component that is ejected from the halo, but also the metal mass in each phase. Metals are produced in stellar evolution. In the model, we adopt the so-called “instantaneous recycling approximation” to treat metal production. For a unit of cold gas mass forming stars, we assume that a fraction of the mass  $R$  is instantaneously returned back to the ISM and a fraction of the mass  $p$  is turned into metal mass and mixed into the ISM,  $\delta M_Z = p \delta M_*$ . Assuming a Chabrier IMF (Chabrier 2003), we adopt the mass return fraction to be  $R = 0.43$  and the total metal yield to be  $p = 0.03$  in this paper. The produced metals are mixed into the cold gas in the disk. The metallicity of the cold gas is simply defined as the ratio of total metal mass in the cold gas and the total cold gas mass,  $Z_{\text{cold}} = M_{Z,\text{cold}}/M_{\text{cold}}$ . New stars have the average metallicity of the cold gas at the time they formed. As star formation proceeds, more metal mass in the galaxy is locked into the stellar mass. Similarly, the stellar metallicity in the model is defined as the ratio of total metal mass in stars and the total stellar mass,  $Z_* = M_{Z,*}/M_*$ . When stellar feedback drives outflow, the metallicity of the outflow is assumed to be equal to the metallicity of the cold gas in the galaxy, which is an assumption widely adopted in SAMs. When the ejected baryons are reincorporated into the halo, the metals are also brought back into the halo and mixed with the halo gas. The model computes the radiative cooling rate of the hot halo gas with a given metallicity by interpolating the cooling tables of Sutherland & Dopita (1993) for varying temperatures and metallicities. We adopted the value for the solar metallicity as  $Z_{\odot} = 0.0134$  (Asplund et al. 2009) when we compare the model predictions with observations.

To summarize, Table 1 provides a brief summary of all of the model parameters in the ejective model and the extended model. We also list the prior ranges over which we allow the parameters to vary.

### 2.3. Constraining Data Sets

In this paper, we explore how the stellar mass function (SMF) and the stellar mass–metallicity relation (MZR) of MW satellite galaxies constrain galaxy formation physics. We use MCMC to explore the parameter space and to sample the

posterior distribution of the model parameters under the data constraints of the MW satellite galaxies.

For the stellar mass function, we adopt the likelihood function defined in Lu et al. (2016) to perform Bayesian inferences. The likelihood function is described by the negative binomial distribution proposed by Boylan-Kolchin et al. (2010) for describing the distribution of the subhalo mass function predicted by  $N$ -body simulations:

$$L(D_{\text{SMF}}|\theta) = \prod_i P(N_i|r, p_i) = \prod_i \frac{\Gamma(N_i + r)}{\Gamma(r)\Gamma(N_i + 1)} p_i^r (1 - p_i)^{N_i}, \quad (11)$$

where  $N_i$  is the observed number of satellite galaxies for a given stellar mass bin  $i$  per MW halo; the two parameters  $r$  and  $p_i$  are determined by the model as  $r = 1/s_1^2$ ,  $p_i = 1/(1 + s_1^2 \mu_i)$ ;  $\mu_i$  is the expected number for the  $i$ th mass bin predicted by the model.  $s_1$  is the fractional scatter from the intrinsic scatter,  $\sigma_{1,i}$ , with respect to the Poisson scatter,  $\mu_i$ , defined as  $s_1 \equiv \sigma_{1,i}/\mu_i$ . We note that the value of  $s_1$  may vary as a function of mass bin and can be simulated for any given model if a large number of merger trees are utilized. In this paper, however, we assume it is a constant  $s_1 = 0.25$ , as we have shown in Lu et al. (2016) that this parameter does not have a strong impact on the likelihood of the data given a predicted stellar mass function. For the observational data, we adopt the stellar masses and memberships of MW satellite galaxies compiled in McConnachie (2012) and only use the properties of the 11 most massive satellite galaxies ( $M_* = 2.9 \times 10^5 M_{\odot}$ ) to constrain the model. Tollerud et al. (2008) showed that the incompleteness of the MW satellite galaxy count becomes important only for fainter dwarfs with  $M_v > -7$  or  $L < 10^5 L_{\odot}$  unless there is a significant low-surface-brightness population of satellites like Crater 2 (Torrealba et al. 2016) with  $L > 10^5 L_{\odot}$ . We restrict this study to only the bright end of the stellar mass function to avoid uncertainties in incompleteness corrections.

We use the observational results of Kirby et al. (2013) as the data constraint on the stellar mass–metallicity relation of MW satellite galaxies. We run the model to predict the stellar-phase metallicity of all satellite galaxies in all of the MW hosts. To fairly compare the model predictions with observations, we only use the satellites with stellar mass higher than  $10^3 M_{\odot}$ . Our resolution tests indicate that our simulations produce converged results for galaxies with stellar mass  $\gtrsim 10^4 M_{\odot}$ . Nevertheless, including galaxies one order of magnitude lower in stellar mass does not affect the overall trend of the stellar mass–metallicity relation predicted by the model over a large range of stellar masses. We then use a second-order polynomial function to fit the relation between the predicted stellar-phase metallicity and stellar mass for all of the simulated samples at  $z = 0$  as

$$\log Z(M_*) = \log Z_0 + c_1 \log M_* + c_2 \log M_*^2, \quad (12)$$

where  $M_*$  is the stellar mass predicted for satellite galaxies, and  $Z$  is the stellar-phase metallicity of a galaxy. The fit is dominated by model galaxies with  $M_* > 10^4 M_{\odot}$ . Tests show that including galaxies slightly below the conservatory resolution limit does not affect the results. The stellar-phase metallicity of a galaxy in the SAM is computed as the fraction

**Table 1**  
Summary of the Semi-analytic Model Parameters

Notation	Meaning of Parameter	Prior	Ejective Model Posterior	Extended Model Posterior
$\Sigma_{\text{SF}}$	Combined parameter for gas-disk size and the gas surface density threshold for star formation	$\log \Sigma_{\text{SF}} \in [-1, 3]$	[1.74, 2.23]	[1.97, 2.88]
$\alpha_{\text{LD}}$	Normalization of the mass-loading factor (Equation (7))	$\log \alpha_{\text{LD}} \in [-12, 2]$	[-4.94, 0.475]	[-10.6, -2.35]
$\beta_{\text{LD}}$	Power-law index for the circular velocity dependence of the mass-loading factor (Equation (7))	$\beta_{\text{LD}} \in [-4, 15]$	[2.33, 7.87]	[-2.33, 6.15]
$V_{\text{out}}$	Characteristic halo circular velocity [ $\text{km s}^{-1}$ ], below which all outflow mass leaves the host halo	$\log V_{\text{out}} \in [0.5, 2.5]$	[1.44, 2.15]	[0.808, 2.15]
$\beta_{\text{out}}$	Steepness of the transition from total outflow for halos with $V_c \ll V_{\text{out}}$ to no outflow for halos with $V_c \gg V_{\text{out}}$	$\beta_{\text{out}} \in [0, 8]$	[1.83, 6.83]	[1.30, 6.70]
$\gamma_{\text{RI}}$	Fraction of outflow mass reincorporated back into the halo (Equation (9))	$\log \gamma_{\text{RI}} \in [-3, 0]$	[-2.95, -2.56]	[-1.68, -0.353]
$M_{\text{pr}}$	Characteristic halo mass scale [ $M_{\odot}$ ] in prevention feedback (Equation (10))	$\log M_{\text{pr}} \in [8, 12]$	...	[10.4, 11.1]
$\beta_{\text{pr}}$	Parameter for the mass dependence in prevention feedback (Equation (10))	$\beta_{\text{pr}} \in [0, 4]$	...	[1.76, 3.59]
$\gamma_{\text{pr}}$	Parameter for the redshift dependence in prevention feedback (Equation (10))	$\gamma_{\text{pr}} \in [0, 4]$	...	[0.389, 1.76]
Comment			Poor fit to SMF	Good fit to SMF and MZR

**Note.** The first column lists the notations of the free parameters in the SAM. The second column briefly explains the meaning of each free parameter. The third column lists the prior for each parameter in our inference. The fourth and the fifth columns list the posterior bounds enclosing 67% marginalized posterior Probability for the ejective feedback model and the extended model, respectively. Both models are constrained to the stellar mass function and the mass–metallicity relation of MW dwarf galaxies. The posterior distributions are shown in Figure 2.

of metal mass that is locked in long-lived stars. The best-fit model is used to represent the mean relation between  $\log \bar{Z}_*$  and  $\log M_*$ , where  $\log \bar{Z}_*$  is the expected logarithmic metallicity given the stellar mass  $M_*$ . When the mean relation is determined for each model, we compute the standard deviation of the predicted model galaxies about the mean relation,  $\sigma_z$ . We assume that the standard deviation is a constant over the stellar mass range probed ( $10^3 M_\odot < M_* < 10^8 M_\odot$ ), which seems a good approximation for the data (Kirby et al. 2013). Kirby et al. (2013) showed that

$$\log \left( \frac{Z_*}{Z_\odot} \right) = 0.3 \pm 0.02 \log \left( \frac{M_*}{10^6 M_\odot} \right) + (-1.69 \pm 0.04), \quad (13)$$

where the best-fit slope  $a_{\text{obs}} = 0.3$  and intercept  $b_{\text{obs}} = 1.69$ , and their  $1\sigma$  errors  $\sigma_a = 0.02$  and  $\sigma_b = 0.04$ . We then assume that the likelihood for a MW satellite galaxy with stellar mass  $M_*$  to have stellar-phase metallicity  $Z_*$  follows a Gaussian function with the mean and the standard deviation determined by the model. Thus, the likelihood function can be written as

$$p(Z_*|M_*) = \exp \left\{ -\frac{[Z_* - \bar{Z}_*(M_*)]^2}{2\sigma_z^2} \right\}. \quad (14)$$

We assume that each observed galaxy is independent from one another when computing the likelihood. Hence, the joint likelihood for a given set of observed galaxies is simply

$$L(D_{\text{MZR}}|\theta) = \prod_i P(Z_{*,i}|M_{*,i}). \quad (15)$$

We note that the data constraint for the mass–metallicity relation extends to much lower masses than the stellar mass function we use because counts of MW dwarfs with  $M_* < 10^5 M_\odot$  likely suffer strongly from incompleteness. In our model for the likelihood function of the mass–metallicity relation, we have assumed that the low-mass satellites detected are representative of the full population at a given mass. This may not be true if, for example, the satellites missed due to incompleteness have systematically different metallicities because the undetected ones are distributed farther away from the MW and have different star formation histories (and thus metallicities) than the detected satellite population. We have tested this by using several randomly selected models with different model parameters. We find that the predicted metallicity of satellite galaxies does not systematically change with the distance between the satellite and the MW halo center. The stellar mass is always the main factor driving the change in metallicity. Therefore, we expect that our approach based on the likelihood function model proposed here should yield robust results.

Finally, we multiply the likelihood for the stellar mass function and the likelihood for the mass–metallicity relation together under the assumption that these two data sets are independent, to arrive at the total likelihood

$$L(D|\theta) = L(D_{\text{SMF}}) \times L(D_{\text{MZR}}). \quad (16)$$

We use MCMC to sample the posterior probability density distribution. In addition to the previous MCMC result adopted

in Lu et al. (2016), two runs for each of the two competing models are presented in this paper. A brief description of the model parameters and priors are listed in Table 1. For a detailed explanation of these parameters, readers are referred to Lu et al. (2014b). For each model, we run MCMC for 20,000 iterations with 144 parallel chains using the differential evolution algorithm (Ter Braak 2006). The convergence test is done with the Gelman–Rubin test (Gelman & Rubin 1992), requiring the potential scale reduction factor  $\hat{R} < 1.2$ . After removing outliers and pre-burn-in states, we obtain  $\sim 650,000$  posterior samples from MCMC for each run.

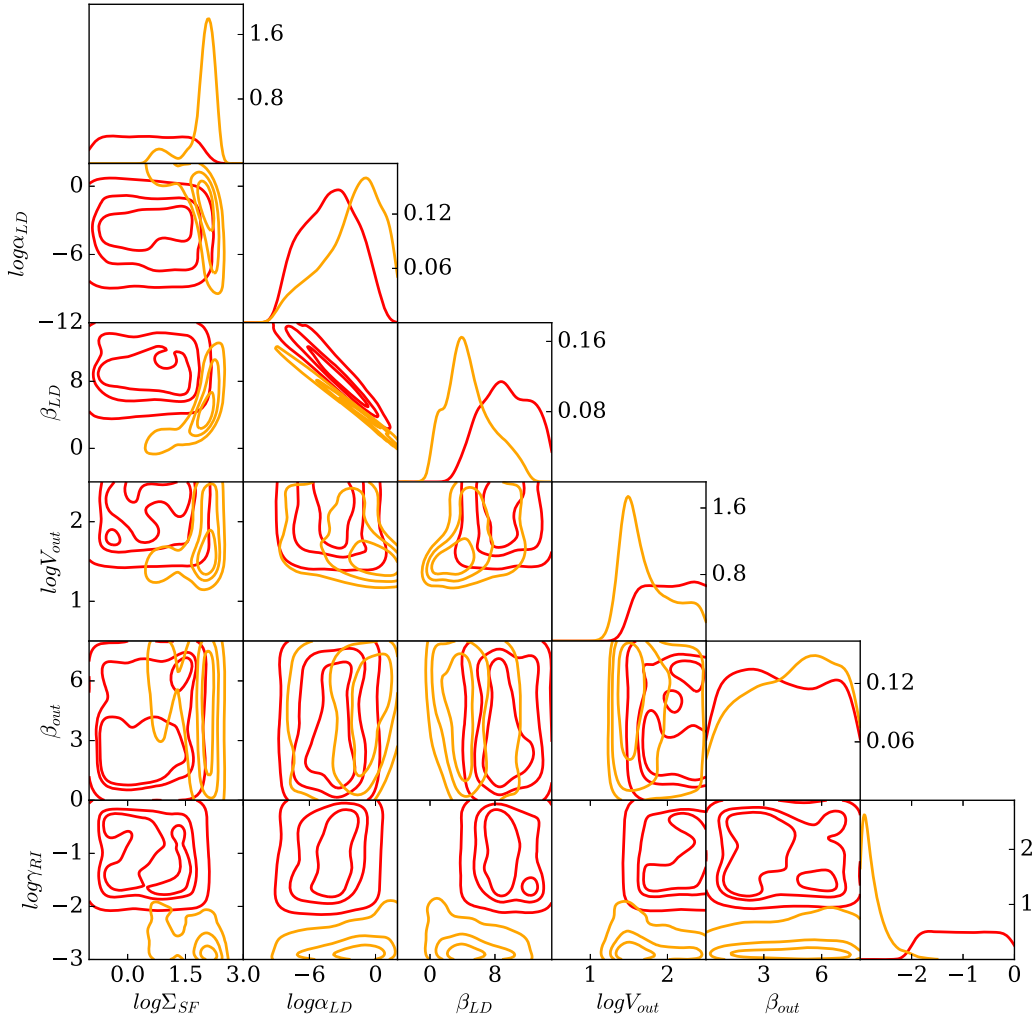
### 3. Results

#### 3.1. Posterior Distribution

We show the posterior probability distributions of the ejective model when it is constrained to different data in Figure 1. The panels on the diagonal line in the figure show the posterior distributions marginalized to each dimension of the parameters, and the off-diagonal panels show the two-dimensional marginalized posterior probabilities. In each panel, the red lines or contours denote the posterior obtained using only the MW SMF as data constraint. The same results have been presented in Lu et al. (2016), but we repeat the plot here to compare with the new results. The yellow lines and contours denote the new results constrained by both the MW SMF and the MZR. As one can see, the addition of the metallicity relation in the data strongly constrains the model. In the previous result with the stellar mass function only, the posterior distributions for many parameters were very broad. When the mass–metallicity relation is added into the constraints, the posterior becomes much narrower for almost all of the free parameters and, more interestingly, some of the modes of the posterior distribution move around in the parameter space. We note that some of the posterior distributions are approaching the prior boundaries, potentially indicating that the prior should be further extended. We have tested varying the priors and found that further extending the priors does not change our conclusions.

We find that the addition of the mass–metallicity relation requires a different outflow mass-loading factor than when matching the stellar mass function alone. The normalization of the outflow mass-loading factor is constrained to higher values, but its halo circular velocity dependence is constrained to be weaker. This is reflected in the location of the modes in the contours of the  $\log \alpha_{\text{LD}}$  versus  $\beta_{\text{LD}}$  panel. From Lu et al. (2016), the stellar mass function-constrained contours have a mode at  $\beta_{\text{LD}} \sim 7$ , but the mass–metallicity relation-constrained contours have a mode at  $\beta_{\text{LD}} \sim 2$ . As Lu et al. (2016) have shown, the rapid increase in the mass-loading factor of the model results in a significantly steeper MZR, underpredicting the stellar-phase metallicity for MW dwarfs. Although the normalization of the mass-loading factor in the newly constrained model is constrained to higher values, the mass-loading factor increases significantly more slowly as the halo circular velocity decreases, resulting in a much better fit to the MZR of the MW satellite galaxies. As we will discuss in the next subsection, however, the model fails to simultaneously fit both the mass–metallicity relation and the stellar mass function even when the parameters are allowed to vary in a broad range of priors.





**Figure 1.** Posterior distributions of the ejective feedback model constrained to different data. The red lines and contours show the one-dimensional and two-dimensional marginalized posterior probability distributions of the ejective model constrained to only the MW satellite stellar mass function. The yellow lines and contours show the same distributions of the same model constrained to both the MW satellite stellar mass function and the stellar-phase metallicity–stellar mass relation of MW satellite galaxies.

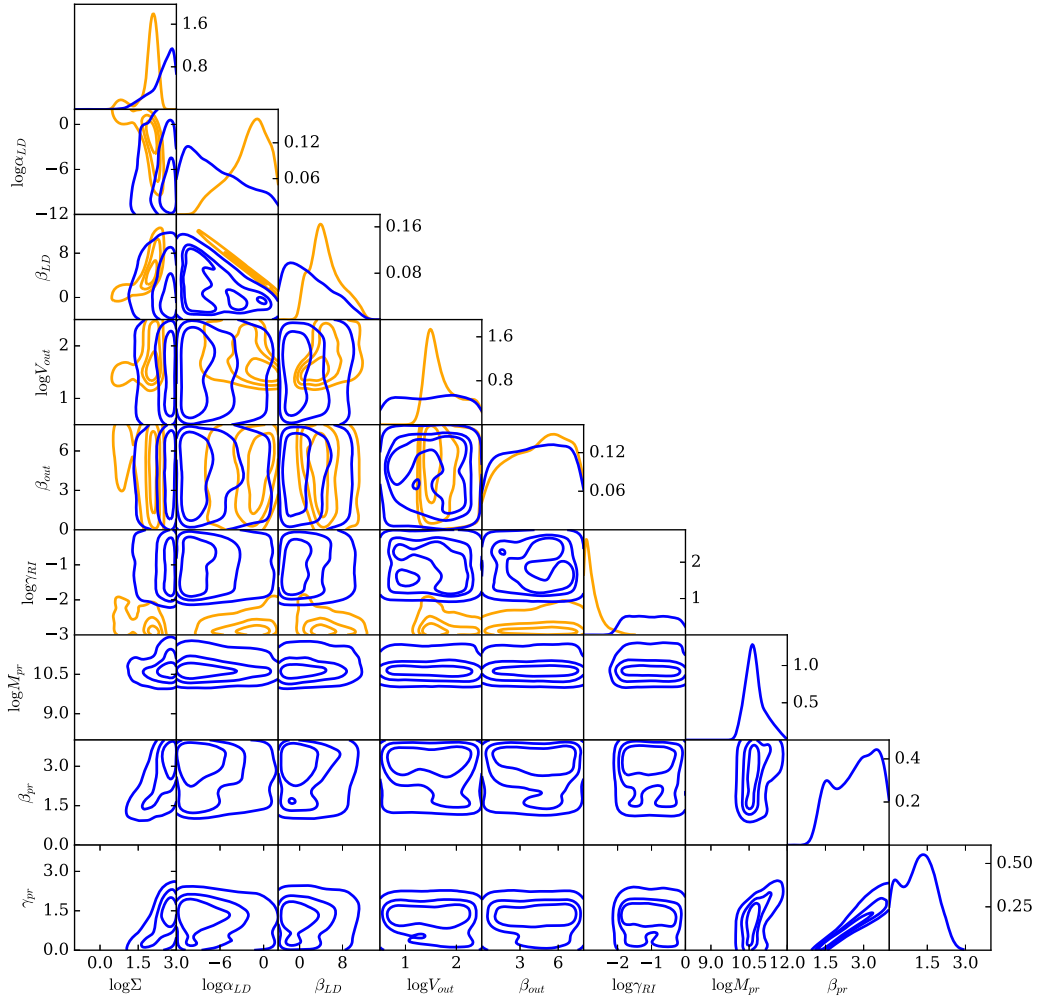
In addition, we find that the posterior distributions of other parameters are also altered to different degrees. The parameter  $\Sigma_{\text{SF}}$  is constrained to slightly higher values, indicating that a higher cold gas surface density or larger gas-disk size is preferred by the addition of the MZR in the data constraints. The parameter  $V_{\text{out}}$  is constrained to slightly lower values. Because halos with circular velocity higher than  $V_{\text{out}}$  retain outflow within the halo, this change indicates that the MZR prefers models that can retain outflow so the metal-enriched gas can fall back on the galaxy. For the gas mass ejected out of the halo, the newly constrained model prefers significantly lower  $\gamma_{\text{RI}}$ , which means that the model is pushed to have very inefficient reincorporation to better match the data. In the next subsection, we will marginalize all the variations of the parameters in the analysis and assess the model in comparison with the data.

Figure 2 shows the posterior distribution of the extended model, which includes both ejective feedback and preventive feedback, constrained to both the MW satellite stellar mass function and the mass–metallicity relation. In comparison, we also include the posterior distribution of the ejective model constrained to the same data sets in the panels where the parameters are common to both models. Since the last three

parameters,  $\log M_{\text{pr}}$ ,  $\beta_{\text{pr}}$ , and  $\gamma_{\text{pr}}$ , are newly introduced in the extended model, there are no counterpart contours and distribution functions for those parameters from the ejective model. The marginalized posterior distribution for the three additional parameters have the following characteristics. First, the characteristic prevention mass scale is strongly constrained to  $M_{\text{pr}} \sim 10^{10.7} M_{\odot}$ , indicating that halos with mass lower than this would accrete baryons at a lower rate than the halo mass accretion rate multiplied by the cosmic baryon fraction at late times. Second, at a fixed redshift, the baryon accretion fraction governed by the preventive feedback scales with the halo mass as  $M_{\text{vir}}^{\beta_{\text{pr}}}$ , where the parameter  $\beta_{\text{pr}}$  is constrained to a fairly broad distribution peaking at  $\sim 3$ . Third, the characteristic prevention mass scale increases with increasing redshift as  $\exp(\gamma_{\text{pr}} z)$ , where the parameter  $\gamma_{\text{pr}}$  is also constrained to a fairly broad distribution peaking at  $\sim 1.4$  in strong degeneracy with varying  $\beta_{\text{pr}}$ . Recall that reionization is included in all of our models as discussed in Section 2.2.1. The fact that the parameters for the preventive feedback model are strongly constrained by the data suggests that preventive feedback in addition to reionization is needed to better fit the data.

In addition to the features for the new parameters in the posterior distribution, we also find that when the preventive





**Figure 2.** Posterior distributions of two different models that are constrained to both the MW satellite stellar mass function and the stellar-phase metallicity–stellar mass relation. The blue lines and contours denote the results of the extended model including both ejective and preventive feedback. The yellow lines and contours are the same as those shown in Figure 1, denoting the ejective model constrained to the same data sets.

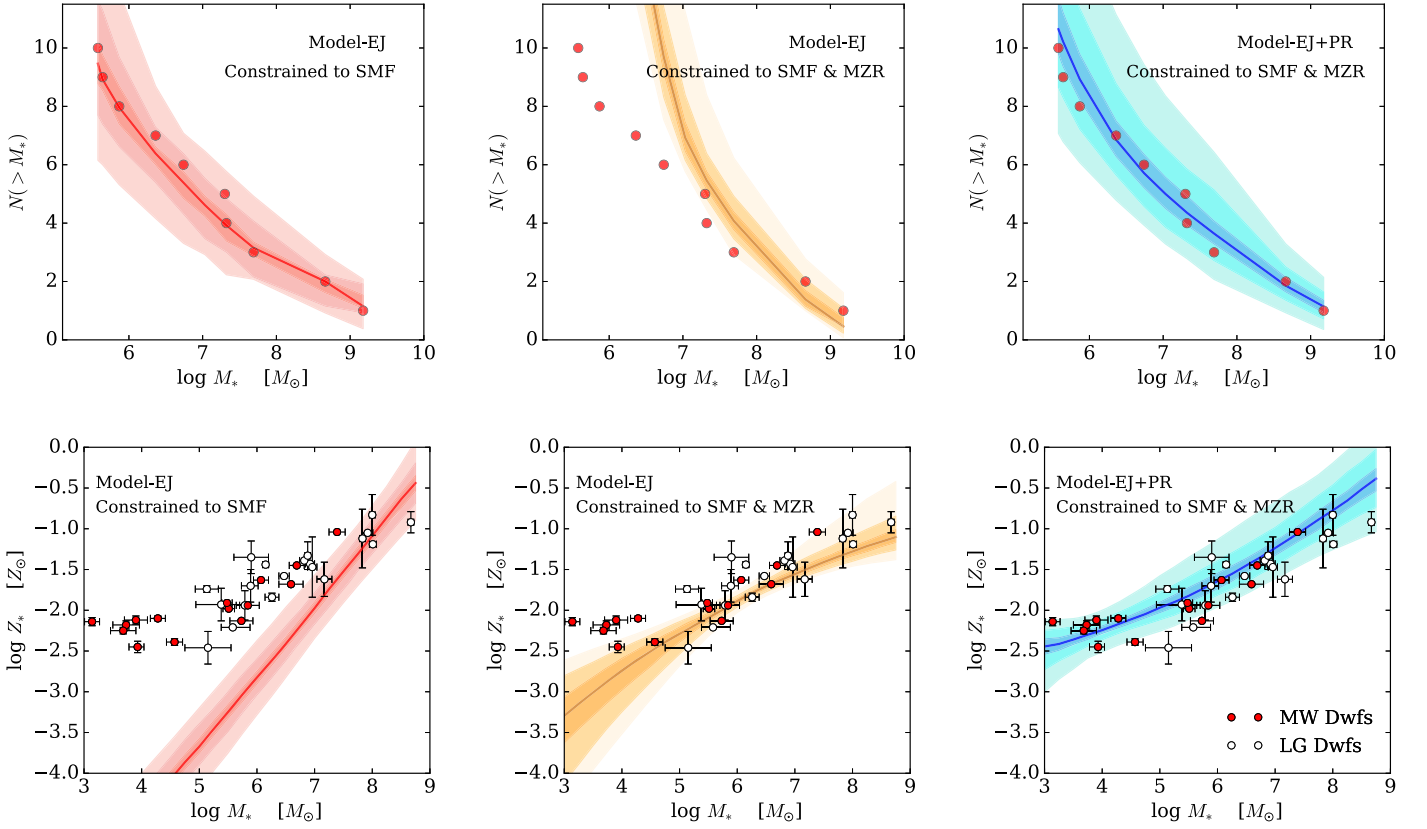
feedback model is included, the parameters for the outflow model are constrained to different values. First of all, the extended model requires weaker outflow to fit the data. This can be found by looking at the marginalized posterior probability distribution of the parameter  $\log \alpha_{LD}$ , which is the normalization of the mass-loading factor. This parameter is required to have high values for the pure ejective model, but is now constrained to have much lower values with a more extended distribution within the prior range, indicating that only a moderate level of outflow is needed to explain the stellar mass function and the mass–metallicity relation. The scaling relation for the outflow mass-loading factor is also constrained to be a weaker function of the halo circular velocity. This can be seen in the posterior probability distribution of parameter  $\beta_{LD}$ . Comparing the posterior distribution of this parameter for the two models, one can find that the peak of the distribution for  $\beta_{LD}$  in the extended model moves to lower values,  $\beta_{LD} \sim 0$ , within a fairly broad range. We find that the “energy-driven wind” model ( $\beta_{LD} = 2$ ) and the “momentum-driven wind” model ( $\beta_{LD} = 1$ ) can explain the mass–metallicity relation fairly well when the other parameters are tuned accordingly. The results suggest that when a certain fraction of baryonic matter is prevented from collapsing into halos in the first place, a large range of possible outflow models with lower

mass-loading factors can reproduce the stellar mass function and the mass–metallicity relation. Moreover, other parameters are also changed accordingly. The parameter  $\Sigma_{SF}$  characterizing the cold gas surface density threshold for star formation and the size of the cold gas disk is constrained to higher values. The transition halo circular velocity,  $V_{out}$ , now has a flat distribution extending to very low values, indicating that the model prefers that outflow gas is retained in the halo rather than leaving the halo. For the outflow gas that leaves the halo, the larger values of  $\gamma_{RI}$  indicate a more rapid reincorporation.

### 3.2. Model Fit

By marginalizing the posterior distribution of the model parameters, we show the predictive distributions for the stellar mass function (upper panels) and the mass–metallicity relation (lower panels) in Figure 3. In each panel, the bands with different intensities of different colors show the 20%, 50%, and 80% predictive distribution.

First, the panels in the left column of Figure 3 show the predictions of the ejective model constrained to only the MW satellite SMF. As discussed in Lu et al. (2016), when the model is constrained to the MW satellite SMF, it systematically predicts a steep mass–metallicity relation and significantly underpredicts metallicity for low-mass satellite galaxies. For



**Figure 3.** Upper panels: posterior predictive distribution of the MW satellite stellar mass function predicted by the three models. The red circles denote the observational data from McConnachie (2012). Lower panels: the posterior predictive distribution of the stellar-phase metallicity as a function of stellar mass of the same models. The circles denote observational data from Kirby et al. (2013). The red filled circles are for the MW dwarfs, and the open circles are for dwarf galaxies in the Local Group but outside the MW. Only the MW dwarfs are used as data constraint in this paper. The left column shows the predictions made by the ejective feedback model that is constrained only to the MW SMF as presented in Lu et al. (2016). The middle column shows the predictions made by the ejective model that is constrained to both the SMF and the MZR of MW dwarfs. Note that the model actually only fits the MZR marginally well, but fails to match the SMF. The right column shows the extended model including both ejective and preventive feedback and fits both the SMF and the MZR of MW dwarfs remarkably well. In each panel, the color bands from dark to light encompass the 20%, 50%, and 80% predictive distribution.

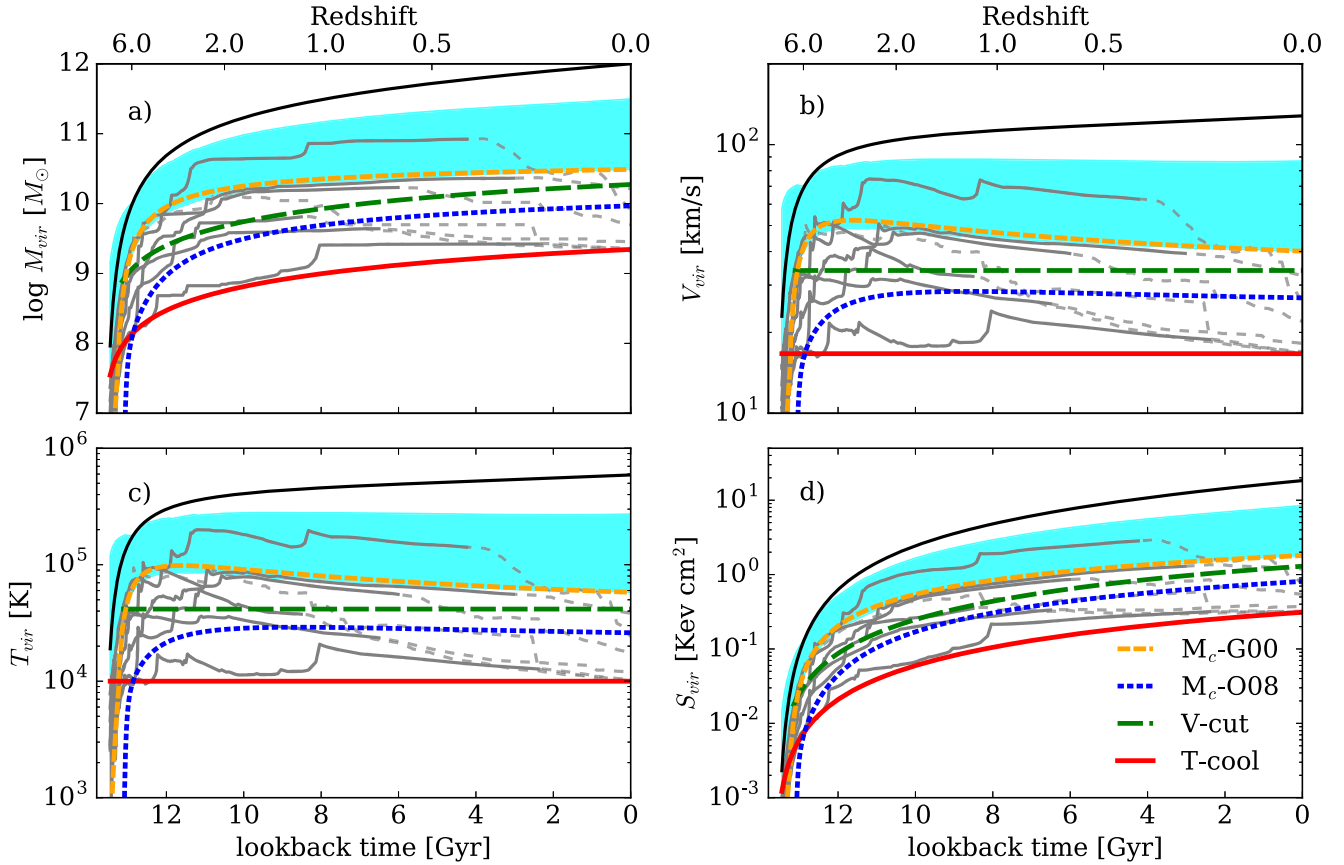
dwarfs with  $M_* < 10^{5.5} M_\odot$ , the model underpredicts the stellar-phase metallicity by a factor of a few. It leaves a question to be answered—can ejective models simultaneously reproduce the SMF and the MZR of the MW satellite galaxies if we constrain the free parameters using both observational data sets?

We answer this question by taking both the SMF and the MZR to constrain the model. MCMC exhaustively explores the parameter space and samples the posterior probability distribution of the free parameters under the constraints of both data sets. By marginalizing the posterior, we produce the predictive distribution for the MW SMF and MZR and find that the model is not able to fit both data sets simultaneously. As we show in the middle column of Figure 3, the model predicts too many satellite galaxies with stellar mass below  $10^7 M_\odot$ , resulting in a satellite SMF significantly steeper than observed. On the other hand, the fit to the mass–metallicity relation is largely improved. For galaxies with  $M_* > 10^5 M_\odot$ , the model recovers the shallow slope of the MZR. For lower stellar masses, however, the model still tends to predict a steeper slope for the MZR. These results suggest that the fit is more strongly influenced by the MZR than the mass function. Nevertheless, it is clear that the ejective model family we explore in this paper cannot simultaneously fit both the mass function and the MZR even when a large range of the parameter space is explored.

Using the result of the flexible model with MCMC, we conclude that the failure of the ejective model is generic but not specific to particular choices of parameter value, and confirm the previous results that additional physics other than outflow rates is needed to resolve the discrepancy in low-mass galaxies (e.g., Font et al. 2011; Henriques et al. 2013; Hou et al. 2016).

Finally, the right column of Figure 3 shows the predictions of the extended model that include both ejective and preventive feedback and is constrained to the MW SMF and MZR. When preventive feedback is included, the model can fit both data sets remarkably well. The constrained model recovers the observed stellar MZR over the entire range of  $M_*$  while simultaneously reproducing the SMF.

We remind the readers that the preventive and ejective forms of feedback are implemented very differently in the model. Preventive feedback reduces the accretion of baryons onto dark-matter halos, and ejective feedback removes a fraction of metal mass from galaxies, resulting in a simultaneous fit to both the SMF and the MZR. The success of the phenomenological model suggests that both modes of feedback are required to accurately model the MW’s satellites. Expulsion of metals via outflows is needed to match the MZR, and suppression of accretion is necessary to reduce the fuel for star formation and reproduce the SMF.



**Figure 4.** Histories of the halo virial mass ( $M_{\text{vir}}$ ; panel a), virial velocity ( $V_{\text{vir}}$ ; panel b), virial temperature ( $T_{\text{vir}}$ ; panel c), and the corresponding virial entropy ( $S_{\text{vir}}$ ; panel d), as functions of lookback time. The gray lines show the seven subhalos randomly selected from all subhalos in our simulations that are more massive than  $2 \times 10^9 M_{\odot}$  at  $z = 0$ . The solid segment of each line denotes the regime where the halo is a distinct halo in the field, and the faded dashed segment denotes the regime when the halo is identified as a subhalo of another halo. The solid black line in each panel shows a smoothed mass-assembly history of a typical MW-size host halo. The red line shows the atomic cooling limit, corresponding to  $T = 10^4$  K. The green dashed line shows the critical accretion limit proposed in Font et al. (2011), i.e.,  $V_{\text{vir}} = 34 \text{ km s}^{-1}$  since  $z = 10$ , to mimic a strong global and local reionization in their model. The blue dotted line shows the fitting function provided by Makiya et al. (2016) for the critical accretion mass affected by reionization as a function of redshift predicted in the hydro simulations of Okamoto et al. (2008). The orange dashed line shows the filtering mass predicted by the reionization model of Gnedin (2000). The cyan band covers 80% of the posterior range of the characteristic prevention mass scale in the preventive feedback model, at which mass scale halos can only accrete half of the cosmic baryon fraction.

### 3.3. Strength of Preventive Feedback

In Figure 4, we demonstrate how the constrained preventive model affects the satellite galaxies. We plot the assembly history of seven subhalos randomly selected from subhalos of the simulated MW host with a final mass  $M_{\text{sub}} > 2 \times 10^9 M_{\odot}$  at  $z = 0$ . We use this random selection to eliminate host-to-host variation. Among the four panels, panel a shows the halo mass-assembly histories; panel b shows the virial velocities; panel c shows the virial temperature; and panel d shows the virial entropy as a function of time (redshift). Each subhalo is denoted by a gray line in the figure, with the solid line denoting the regime where the halo is a distinct halo in the field and the faded dashed line denoting the regime since the halo is accreted for the first time into another halo. The virial velocity is defined as the halo circular velocity at the virial radius. The virial temperature of a halo is related to the circular velocity as  $T_{\text{vir}} = 36.7 (V_{\text{c}}/\text{km s}^{-1})^2 \text{ K}$ . Following Lu et al. (2015b), we define the virial entropy as

$$S_{\text{vir}} = \frac{T_{\text{vir}}}{n_{\text{vir}}^{2/3}}, \quad (17)$$

with  $T_{\text{vir}}$  the virial temperature of the halo, and  $n_{\text{vir}}$  the mean gas particle number density of a virialized halo assuming the

cosmic baryon fraction  $f_{\text{b},0}$ . These quantities have a meaningful impact on halos only when they are distinct halos (represented by the solid part of the gray lines in Figure 4). When the halo becomes a subhalo, its mass and radius take the values assigned by the ROCKSTAR halo identifier (Behroozi et al. 2013). We still use these values to compute the “virial velocity,” the “virial temperature” and the “virial entropy,” but they no longer have the same physical meaning as for distinct halos. We also note that after a halo becomes a subhalo, it is not subject to preventive feedback but completely loses its halo gas. Although it stops accreting new gas, star formation can still continue until the existing cold gas is exhausted. The figure shows that the masses of those halos increase rapidly in the first  $\sim 2$  Gyr of the universe (before  $z \sim 2$ ). At late times, most of the subhalos lose their mass due to tidal stripping as they are affected by the tidal field of the MW host. As one can see, MW subhalos that host classical dwarfs are accreted into another halo typically between  $z = 1.5$  and  $0.5$ , consistent with the results of Wetzel et al. (2015).

We also show predictions from a handful of models that affect galaxy formation by preventing baryons from cooling or collapsing into halos. The red line in each panel shows the

atomic cooling limit, corresponding to  $T = 10^4$  K (red line). Halos with a virial temperature below the atomic cooling limit are not able to cool baryons to form a galaxy in its potential well unless positive feedback of reionization is considered (for instance, ionizing UV radiation promoting  $H_2$  formation; see Ricotti et al. 2002; Bovill & Ricotti 2009). As we can see in the figure, the atomic cooling limit only affects halos when their circular velocities are lower than  $17 \text{ km s}^{-1}$ . It is clear that the atomic cooling limit itself is not sufficient as a preventive process to affect most of the observed classical dwarfs.

The other process is the heating due to UV photoionization (photoheating, for short), which not only offsets cooling losses, but also prevents low-mass halos from accreting their full complement of baryons. We show the characteristic mass scales in two photoheating models. First, we show the characteristic halo mass introduced in the model proposed by Font et al. (2011). In the model, the authors proposed a model in which halos with  $V_c \leq 34 \text{ km s}^{-1}$  cannot cool baryons due to the combination of global reionization and the photoheating in the Local Group for  $z < 10$ . We show this scale by a green dashed line in each panel of Figure 4. This scale is higher than the atomic cooling scale but has a similar behavior, as it corresponds to a constant virial temperature threshold of  $4 \times 10^4$  K. The second photoheating model we consider is the “filtering mass” proposed by Gnedin (2000) based on linear perturbation theory (orange dashed lines in Figure 4). The model predicts that halos with masses lower than some characteristic mass (the filtering mass) accrete baryons at a significantly reduced rate. The filtering mass is higher than in the Font et al. (2011) model, predicting a stronger prevention for halos with higher masses. In the Gnedin (2000) model, the baryon fraction scales with halos as  $f_b \propto M^3$ , while the Font et al. (2011) model imposes a sharp truncation. Therefore, the Font et al. (2011) model has a stronger effect on halos with lower masses. Moreover, Okamoto et al. (2008) derived a model for the suppression of baryonic accretion into halos using hydrodynamical simulations. As shown in Figure 4, the critical mass scales predicted by the simulations are considerably lower than the model adopted here based on Gnedin (2000), suggesting that considering the Okamoto et al. (2008) model will not affect the preventive model studied in this paper.

Lastly, the cyan band in Figure 4 shows the preventive model constrained by the MW satellite galaxy SMF and the MZR in this work. The bands show the characteristic halo mass, circular velocity, temperature, and virial entropy at which only half of the baryon mass can collapse into the halo as a function of redshift. The band covers 80% of the posterior distribution of the constrained model. Below this characteristic halo mass, the baryon fraction in the halos scales as  $f_b \propto M_{\text{vir}}^{\beta_{\text{pr}}}$ , as described in Equation (10). In general, the model requires that the characteristic prevention mass scale increases with time. At early times ( $z > 6$ ), the mass scale can be lower than  $10^9 M_\odot$ , and becomes higher ( $> 10^{10} M_\odot$ ) at late times. The corresponding halo circular velocity is above  $40 \text{ km s}^{-1}$ , and the virial entropy is a few  $\text{KeV cm}^2$ . With this level of prevention, most of the MW subhalos, except a couple of the most massive ones, are generally affected by preventive feedback. Interestingly, this level of entropy is consistent with the level needed to match the star formation histories of field galaxies with mass equal to or lower than the MW (Lu et al. 2015b). In addition to the normalization, the data require

preventive feedback that varies with redshift in a similar way as the mass-assembly history of the MW host halos (shown as the black solid line in Figure 4), suggesting that the evolution of the strength of preventive feedback may be related to the formation of the host halo or its central galaxy.

#### 4. Discussion and Conclusions

In this paper, we explored the extensive parameter space of a flexible semi-analytic galaxy formation model within the cosmological evolution of  $\Lambda$ CDM dark-matter halos. We have demonstrated that the ejective feedback model, where feedback is captured only in the form of strong outflows, fails to recover the SMF and the MZR of MW dwarf galaxies simultaneously. The strong outflows required to suppress star formation in low-mass halos expel too much mass in the form of metals from low-mass galaxies, resulting a steep mass-metallicity relation that underpredicts the stellar-phase metallicity of MW dwarf galaxies.

This result is in agreement with the hydrodynamic simulations of Torrey et al. (2014), who pointed out that low-mass galaxies need to retain a large fraction of metals produced in star formation to avoid predicting an MZR that was too steep relative to observations. Recent high-resolution hydrodynamic simulations can also achieve remarkable agreement with the observed MZR over a large stellar mass range when a considerable fraction of metals are retained in the inner halos (Anglés-Alcázar et al. 2016; Christensen et al. 2016; Ma et al. 2016; Muratov et al. 2017). On the other hand, our inferences show that adopting a smaller mass-loading factor for low-mass halos, so that the galaxies can retain enough metal mass to match the MZR, results in baryonic mass fractions in low-mass subhalos that are too high, thereby predicting a significantly steeper stellar mass function than is observed. In summary, our inference shows that the ejective feedback model with metal-enriched outflow fails to match the two primary data sets even when many relevant parameters are allowed to vary in a large parameter space. If outflow is the primary feedback that suppresses baryon mass in low-mass halos, the net outflow has to be metal deficient.

The difficulty in fitting the data forced us to implement a preventive feedback model in the SAM. In this model, a fraction of baryons associated with the dark matter that constitutes the gravitational potential of the dwarf galaxies is prevented from collapsing into the low-mass halos. The prevention is increasingly important for lower-mass halos, resulting in a decreasing baryonic-to-dark-matter mass ratio for lower halo masses to match the stellar mass function of dwarf galaxies. At the same time, moderate ejective feedback is still needed to expel a fraction of metal mass that is mixed into the ISM from galaxies to match the MZR. The model with the combination of both preventive and ejective feedback can simultaneously match the SMF and MZR of MW dwarf galaxies. The success of the combined model, coupled with the failure of each model individually, suggests that two different feedback mechanisms are needed to explain the two pieces of data. Specifically, moderate outflow is needed to match the MZR, and a strong preventive feedback is responsible for governing the low star formation efficiency in low-mass halos.

This result is in agreement with a number of previous studies. Using different SAMs, Font et al. (2011) and Hou et al. (2014, 2016) found that lowering the mass-loading factor results in an increased metallicity (also see Guo et al. 2016) and



enhanced reionization, which works in the same way as the preventive model we explored in this paper, is needed to suppress star mass in the MW subhalos. A similar conclusion about the effect of lowering the mass-loading factor on galaxy metallicity is also reached in hydrodynamic simulation studies (e.g., Crain et al. 2015). More recently, using high-resolution hydrodynamical simulations, Christensen et al. (2016) showed that preventive feedback significantly suppressed baryon accretion in low-mass halos. We also note that strong prevention leaves room for metal-enriched outflow (e.g., Dalcanton 2007) to be consistent with the observational constraints.

The comparison between the constrained preventive model and existing reionization models suggests that a higher degree of prevention than normal reionization, similar to the Gnedin (2000) filtering mass implementation, is needed to fit the data. Benson et al. (2002b) also found that global reionization implemented using the Gnedin (2000) formalism has a relatively mild effect. Similarly, using a parameter sensitivity analysis on an SAM, Gómez et al. (2014) found that a stronger prevention than typical reionization models can provide is needed to match the metallicity function of MW satellite galaxies. In the study, the authors found that only  $\sim 0.05$  of the cosmic baryon fraction should be allowed to accrete into halos with circular velocity between  $30 \text{ km s}^{-1}$  and  $50 \text{ km s}^{-1}$  in order to match observations. If the required prevention suggested in the present paper is due to photoheating of local reionization, the stronger prevention in the Local Group than the global reionization would indicate the importance of considering inhomogeneous reionization (e.g., Busha et al. 2010; Lunnan et al. 2012; Li et al. 2014).

We note that there are still large uncertainties in the ejective model, and variations in its prescriptions might provide a better match to the data. One such variation is to assume that metals are differentially expelled from galaxies or preferentially reincorporated into galaxies with respect to hydrogen (i.e., metal-deficient outflows or metal-enriched inflows). For example, if the outflow had lower metallicity than the ISM, it would be possible for the ejective model to predict higher metallicities for low-mass galaxies to match the mass-metallicity relation. Muratov et al. (2017) recently showed that galaxies in hydrodynamic simulations retain a large fraction of metals produced in star formation within their host halos. They found that metal-poor gas was ejected from the ISM due to entrainment in outflows generated by the low-intensity star formation. Alternatively, one can assume that the ejected metals are preferentially reincorporated back into galaxies from the circumgalactic medium (CGM). Li et al. (2010) essentially adopt this type of idea in their SAM by adopting a route to recycle most of the metals produced in newly formed stars through the hot phase. In their model, 95% of newly produced metals are deposited directly into the hot gas for galaxies with a dark-matter halo virial mass less than  $5 \times 10^{10} M_{\odot}$ . The metals then eventually fall back to the galaxy and fuel the next episode of star formation, resulting in a shallow metallicity–luminosity relation.

Furthermore, the ejective feedback has other uncertainties that can affect predictions for low-mass galaxies. Henriques et al. (2013, 2015) showed that when the outflowed mass has a long delay before being reincorporated into low-mass halos, the model can match the evolution of the field galaxy luminosity function remarkably well. Hou et al. (2016) demonstrated a

model where the mass-loading factor decreases with time, as suggested by a recent hydrodynamic simulation (Muratov et al. 2015). Together with saturated outflows at the low-mass end and a strong local reionization, which are effectively similar to our extended model, the model (named “EvoFb-LR” in Hou et al. 2016) can achieve a reasonable fit to both the SMF and MZR of MW dwarf galaxies. While these prescriptions for outflow and reincorporation have been tested by recent hydrodynamical simulations, the actual behavior of the outflow is still a matter of active debate (e.g., Anglés-Alcázar et al. 2016; Christensen et al. 2016).

We note the important caveat that tidal effects from the MW-mass stellar disk can reduce the number of surviving subhalos (e.g., Zolotov et al. 2012; Wetzel et al. 2016; Zhu et al. 2016), which can affect how the ejective model fits the stellar mass function. We stress, however, that the MZR is independent from the satellite abundance. Therefore, the requirement for weaker outflow implied by this relation is robust, even if a fraction of subhalos are destroyed by the MW disk. Still, it is clear that there is some degeneracy between the preventive feedback and the destruction of subhalos. To break this degeneracy, detailed modeling of these processes and quantitative inferences from dwarf galaxies both within and outside the MW will be essential.









Although the model presented here with the addition of preventive feedback can match the data, the physical mechanisms that produce the prevention remain unclear. We have taken an extreme model where baryons are prevented from collapsing into low-mass halos, but the assumption has not been fully tested against observation. At the current stage, other alternative hypotheses, where baryons still make it into the galaxy but are prevented from forming stars by large turbulent pressure or photoheating (e.g., Forbes et al. 2016), can produce similar effects in galaxy formation models. The data constraints adopted in this paper are not able to distinguish between these scenarios even if models for the relevant physics were implemented in the model. Data on the gas and metal content of the ISM and CGM can more directly test whether baryons are accreted into the halo or the galaxy. Future studies using field low-mass galaxies with cold gas measurements will be needed to further test these models.

In summary, a model including reionization and the ejective feedback of gas with well-mixed metals cannot reproduce both the SMF and MZR. However, as we discuss, the parameter space for galaxy formation models is currently poorly constrained, pointing to significant opportunities for future advances. Although the metallicity of low-mass galaxies provides a useful constraint for the model, further tests for different feedback processes and metal recycling processes using broader data constraints are needed. While models in this work are constrained to local data only, we find that preventive feedback needs to have (largely uncertain) redshift dependence. This can be possibly better constrained using the star formation histories of the MW dwarfs (e.g., Weisz et al. 2014). We defer this investigation to a future study, which will help us better understand the origin of the prevention. Moreover, the evolution and the gas content of field galaxies with higher masses ( $M_* > 10^9 M_{\odot}$ ) could potentially help break some of the parameter degeneracies by providing independent constraints on the buildup of stellar mass from high redshift to the present day. Comprehensive model inferences from compilations

of data of field galaxies and dwarf galaxies in the Local Group will improve our understanding of galaxy formation.

We acknowledge the Ahmanson Foundation for providing computational resources used in this work. The zoom-in simulations were performed using computational resources of SLAC National Accelerator Laboratory and of the National Energy Research Scientific Computing Center. We thank the SLAC computational team for their consistent support. Support for programs HST-AR-13896, HST-AR-13888, HST-AR-13270, HST-AR-12836, and HST-GO-14734 was provided by NASA through a grant from the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. We are grateful to Takashi Okamoto for providing tabulated data for the critical mass as a function of redshift presented in Okamoto et al. (2008). Y.L. thanks Romeel Davé, Phil Hopkins, Evan Kirby, Houjun Mo, Josh Simon, and Qingjuan Yu for helpful discussion. A.W. was supported by a Caltech-Carnegie Fellowship, in part through the Moore Center for Theoretical Cosmology and Physics at Caltech. Y.Y.M. is supported by the Samuel P. Langley PITT PACC Postdoctoral Fellowship, and was supported by the Weiland Family Stanford Graduate Fellowship. S.T. was supported by the Alvin E. Nashman Fellowship. M.B.K. acknowledges support from NSF grant AST-1517226 in addition to the *HST* grants listed above.

### ORCID iDs

Yu Lu  <https://orcid.org/0000-0003-2691-1622>  
 Andrew Benson  <https://orcid.org/0000-0001-5501-6008>  
 Andrew Wetzel  <https://orcid.org/0000-0003-0603-8942>  
 Yao-Yuan Mao  <https://orcid.org/0000-0002-1200-0820>  
 Stephanie Tonnesen  <https://orcid.org/0000-0002-8710-9206>  
 Annika H. G. Peter  <https://orcid.org/0000-0002-8040-6785>  
 Michael Boylan-Kolchin  <https://orcid.org/0000-0002-9604-343X>  
 Risa H. Wechsler  <https://orcid.org/0000-0003-2229-011X>

### References

- Anglés-Alcázar, D., Faucher-Giguère, C.-A., Kereš, D., et al. 2017, *MNRAS*, **470**, 4698
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, **47**, 481
- Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013, *ApJ*, **762**, 109
- Benson, A. J. 2010, *PhR*, **495**, 33
- Benson, A. J. 2014, *MNRAS*, **444**, 2599
- Benson, A. J., Frenk, C. S., Lacey, C. G., Baugh, C. M., & Cole, S. 2002a, *MNRAS*, **333**, 177
- Benson, A. J., Lacey, C. G., Baugh, C. M., Cole, S., & Frenk, C. S. 2002b, *MNRAS*, **333**, 156
- Bigiel, F., Leroy, A., Walter, F., et al. 2008, *AJ*, **136**, 2846
- Bovill, M. S., & Ricotti, M. 2009, *ApJ*, **693**, 1859
- Bower, R. G., Vernon, I., Goldstein, M., et al. 2010, *MNRAS*, **407**, 2017
- Boylan-Kolchin, M., Springel, V., White, S. D. M., & Jenkins, A. 2010, *MNRAS*, **406**, 896
- Bryan, G. L., & Norman, M. L. 1998, *ApJ*, **495**, 80
- Bullock, J. S., Kravtsov, A. V., & Weinberg, D. H. 2000, *ApJ*, **539**, 517
- Busha, M. T., Alvarez, M. A., Wechsler, R. H., Abel, T., & Strigari, L. E. 2010, *ApJ*, **710**, 408
- Cautun, M., Frenk, C. S., van de Weygaert, R., Hellwing, W. A., & Jones, B. J. T. 2014, *MNRAS*, **445**, 2049
- Chabrier, G. 2003, *PASP*, **115**, 763
- Chau, A., Mayer, L., & Governato, F. 2017, *ApJ*, **845**, 17
- Christensen, C. R., Davé, R., Governato, F., et al. 2016, *ApJ*, **824**, 57
- Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, *MNRAS*, **319**, 168
- Couchman, H. M. P., & Rees, M. J. 1986, *MNRAS*, **221**, 53
- Cousin, M., Buat, V., Boissier, S., et al. 2016, *A&A*, **589**, A109
- Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, *MNRAS*, **450**, 1937
- Dalcanton, J. J. 2007, *ApJ*, **658**, 941
- Dekel, A., & Silk, J. 1986, *ApJ*, **303**, 39
- Eadie, G. M., Harris, W. E., & Widrow, L. M. 2015, *ApJ*, **806**, 54
- Efstathiou, G. 1992, *MNRAS*, **256**, 43P
- Font, A. S., Benson, A. J., Bower, R. G., et al. 2011, *MNRAS*, **417**, 1260
- Forbes, J. C., Krumholz, M. R., Goldbaum, N. J., & Dekel, A. 2016, *Natur*, **535**, 523
- Gelman, A., & Rubin, D. 1992, *StaSc*, **7**, 457
- Gnedin, N. Y. 2000, *ApJ*, **542**, 535
- Gómez, F. A., Coleman-Smith, C. E., O'Shea, B. W., Tumlinson, J., & Wolpert, R. L. 2012, *ApJ*, **760**, 112
- Gómez, F. A., Coleman-Smith, C. E., O'Shea, B. W., Tumlinson, J., & Wolpert, R. L. 2014, *ApJ*, **787**, 20
- González, R. E., Kravtsov, A. V., & Gnedin, N. Y. 2013, *ApJ*, **770**, 96
- Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, *MNRAS*, **413**, 101
- Guo, Y., Koo, D. C., Lu, Y., et al. 2016, *ApJ*, **822**, 103
- Henriques, B. M. B., Thomas, P. A., Oliver, S., & Roseboom, I. 2009, *MNRAS*, **396**, 535
- Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2013, *MNRAS*, **431**, 3373
- Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, *MNRAS*, **451**, 2663
- Hirschmann, M., De Lucia, G., & Fontanot, F. 2016, *MNRAS*, **461**, 1760
- Hoefl, M., Yepes, G., Gottlöber, S., & Springel, V. 2006, *MNRAS*, **371**, 401
- Hou, J., Frenk, C. S., Lacey, C. G., & Bose, S. 2016, *MNRAS*, **463**, 1224
- Hou, J., Yu, Q., & Lu, Y. 2014, *ApJ*, **791**, 8
- Kauffmann, G., Li, C., Zhang, W., & Weinmann, S. 2013, *MNRAS*, **430**, 1447
- Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, *MNRAS*, **264**, 201
- Kennicutt, J. R. C. 1998, *ApJ*, **498**, 541
- Kennicutt, J. R. C., Calzetti, D., Walter, F., et al. 2007, *ApJ*, **671**, 333
- Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, *ApJ*, **779**, 102
- Kirby, E. N., Lanfranchi, G. A., Simon, J. D., Cohen, J. G., & Guhathakurta, P. 2011, *ApJ*, **727**, 78
- Komatsu, E., Dunkley, J., Nolte, M. R., et al. 2009, *ApJS*, **180**, 330
- Kravtsov, A. V., Gnedin, O. Y., & Klypin, A. A. 2004, *ApJ*, **609**, 482
- Li, T. Y., Alvarez, M. A., Wechsler, R. H., & Abel, T. 2014, *ApJ*, **785**, 134
- Li, Y.-S., De Lucia, G., & Helmi, A. 2010, *MNRAS*, **401**, 2036
- Lu, Y., Benson, A., Mao, Y.-Y., et al. 2016, *ApJ*, **830**, 59
- Lu, Y., Blanc, G. A., & Benson, A. 2015a, *ApJ*, **808**, 129
- Lu, Y., & Mo, H. J. 2007, *MNRAS*, **377**, 617
- Lu, Y., Mo, H. J., Lu, Z., Katz, N., & Weinberg, M. D. 2014a, *MNRAS*, **443**, 1252
- Lu, Y., Mo, H. J., & Wechsler, R. H. 2015b, *MNRAS*, **446**, 1907
- Lu, Y., Mo, H. J., Weinberg, M. D., & Katz, N. 2011, *MNRAS*, **416**, 1949
- Lu, Y., Wechsler, R. H., Somerville, R. S., et al. 2014b, *ApJ*, **795**, 123
- Lunnan, R., Vogelsberger, M., Frebel, A., et al. 2012, *ApJ*, **746**, 109
- Ma, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2016, *MNRAS*, **456**, 2140
- Macciò, A. V., & Fontanot, F. 2010, *MNRAS*, **404**, L16
- Macciò, A. V., Kang, X., Fontanot, F., et al. 2010, *MNRAS*, **402**, 1995
- Makiya, R., Enoki, M., Ishiyama, T., et al. 2016, *PASJ*, **68**, 25
- Mao, Y.-Y., Williamson, M., & Wechsler, R. H. 2015, *ApJ*, **810**, 21
- Mateo, M. L. 1998, *ARA&A*, **36**, 435
- McConnachie, A. W. 2012, *AJ*, **144**, 4
- Mo, H. J., & Mao, S. 2002, *MNRAS*, **333**, 768
- Mo, H. J., & Miralda-Escude, J. 1996, *ApJ*, **469**, 589
- Mo, H. J., Yang, X., van den Bosch, F. C., & Katz, N. 2005, *MNRAS*, **363**, 1155
- Muñoz, R. R., Carlin, J. L., Frinchaboy, P. M., et al. 2006, *ApJL*, **650**, L51
- Muratov, A. L., Kereš, D., Faucher-Giguère, C.-A., et al. 2015, *MNRAS*, **454**, 2691
- Muratov, A. L., Kereš, D., Faucher-Giguère, C.-A., et al. 2017, *MNRAS*, **468**, 4170
- Noh, Y., & McQuinn, M. 2014, *MNRAS*, **444**, 503
- Okamoto, T., Frenk, C. S., Jenkins, A., & Theuns, T. 2010, *MNRAS*, **406**, 208
- Okamoto, T., Gao, L., & Theuns, T. 2008, *MNRAS*, **390**, 920
- Oppenheimer, B. D., Davé, R., Kereš, D., et al. 2010, *MNRAS*, **406**, 2325
- Parry, O. H., Eke, V. R., Frenk, C. S., & Okamoto, T. 2012, *MNRAS*, **419**, 3304
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, **594**, A13
- Pontzen, A., & Governato, F. 2012, *MNRAS*, **421**, 3464
- Ricotti, M., Gnedin, N. Y., & Shull, J. M. 2002, *ApJ*, **575**, 49

- Sawala, T., Frenk, C. S., Fattahi, A., et al. 2016, *MNRAS*, **457**, 1931
- Simon, J. D., & Geha, M. 2007, *ApJ*, **670**, 313
- Somerville, R. S. 2002, *ApJL*, **572**, L23
- Somerville, R. S., & Davé, R. 2015, *ARA&A*, **53**, 51
- Starkenburg, E., Helmi, A., De Lucia, G., et al. 2013, *MNRAS*, **429**, 725
- Sutherland, R. S., & Dopita, M. A. 1993, *ApJS*, **88**, 253
- Ter Braak, C. J. F. 2006, *Stat. Comput.*, **16**, 239
- Thoul, A. A., & Weinberg, D. H. 1996, *ApJ*, **465**, 608
- Tollerud, E. J., Bullock, J. S., Strigari, L. E., & Willman, B. 2008, *ApJ*, **688**, 277
- Torrealba, G., Koposov, S. E., Belokurov, V., & Irwin, M. 2016, *MNRAS*, **459**, 2370
- Torrey, P., Vogelsberger, M., Genel, S., et al. 2014, *MNRAS*, **438**, 1985
- Weisz, D. R., Dolphin, A. E., Skillman, E. D., et al. 2014, *ApJ*, **789**, 147
- Wetzel, A. R., Deason, A. J., & Garrison-Kimmel, S. 2015, *ApJ*, **807**, 49
- Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, *ApJL*, **827**, L23
- White, C. E., Somerville, R. S., & Ferguson, H. C. 2015, *ApJ*, **799**, 201
- White, S. D. M., & Rees, M. J. 1978, *MNRAS*, **183**, 341
- Xue, X. X., Rix, H. W., Zhao, G., et al. 2008, *ApJ*, **684**, 1143
- Zhu, Q., Marinacci, F., Maji, M., et al. 2016, *MNRAS*, **458**, 1559
- Zolotov, A., Brooks, A. M., Willman, B., et al. 2012, *ApJ*, **761**, 71